

Übung zur Lehrveranstaltung  
**Semantic Web Technologies 1**  
Sebastian Rudolph und Duc Thanh Tran  
Wintersemester 2012/13  
<http://semantic-web-grundlagen.de>  
Übung 4: Linked Data and Semantic Search

**Aufgabe 4.1:** Entscheiden Sie, ob die folgenden Aussagen wahr oder falsch sind:

- (a) URIs, IRIs und IDs sind gleichartige Konzepte die für beliebige Ressourcen stehen können.
- (b) URIs und URLs sind unterschiedliche Konzepte. Letzteres wird aber auch verwendet um Ressourcen im Web zu identifizieren.
- (c) Zwei URIs bezeichnen die gleiche Ressource, wenn Sie über owl:sameAs verbunden sind.
- (d) Zwei Ressourcen können nur über owl:sameAs verlinkt werden.
- (e) Antwort auf einen HTTP-Request ist immer in HTML-Format.
- (f) Linked Data ist eine Untermenge von RDF Data.

**Aufgabe 4.2:** Wissensfragen

- Warum wird Linked Data verlinkt?
- Welche Links sollten zur Verlinkung verwendet werden
- Wir wollen Linked Data für KIT erstellen. Wie sieht eine Linked Data Beschreibung für KIT aus, die alle Linked Data Principles erfüllt?
- Sind Webseite die über Hyperlinks verknüpft sind Linked Data? Warum, Warum nicht?
- Können Webseiten die über Hyperlinks verknüpft sind, in Linked Data umgewandelt werden? Wie?
- Was sind Benefits von Linked Data? Auf welche Prinzipien basierend diese Benefits?
- Welches RDF Storage ist gut geeignet für diese Anfragen? Warum?
- Gegeben diese Anfrage:
- Was sind Unterschiede bzgl. Retrieval & Combine bei Keyword Query auf textuelle Daten und Strukturierte Query auf strukturierte Daten
- Was sind Unterschiede bzgl. Retrieval & Combine bei Keyword Search auf textuelle Daten und Keyword Search auf strukturierte Daten

**Aufgabe 4.3:** Rechenaufgabe

- Gegeben 1 Query, 2 Dokumente, und Ihre Terme
- Berechnet Weight für die Terme...

$$w_{t,d} = \frac{tf}{|d|} * idf$$

- Berechnet die Distanz zwischen Query und Dokument

$$Sim(q, d) = Cos((w_{1,d}, ..., w_{t,d}), (w_{1,q}, ..., w_{k,q}))$$

- Rank query auf Basis dieser Distanz
- Gegeben ein Model zur Kombination von Features: Score:  $t * (PageRank) + (1-t) * (Distanz)$ .  
Gegeben Dokument 1 hat PageRank 100 und Dokument 2 mit PageRank 80. Welches Dokument wird nun höher gerankt?

Bob is a good friend of mine. We went to the same university, and also shared an apartment in Berlin in 2008. The trouble with Bob is that he takes much better photos than I do:

**Doc 1**

Bob went to a university in Berlin. We both are good friends. He takes very good photos as an amateur though his camera is not better than mine.

**Doc 2**